Neuro-Oncology
22(3), 311–312, 2020 | doi:10.1093/neuonc/noaa002 | Advance Access date 6 January 2020

Can my computer tell me if this tumor is IDH mutated?

Timothy J. Kaufmann and Bradley J. Erickson

Department of Radiology, Mayo Clinic, Rochester, Minnesota (T.J.K., B.J.E.)

Corresponding Author: Timothy J. Kaufmann, Professor of Radiology, Mayo Clinic, Rochester, MN, 200 1st St. SW, Rochester, MN 55905 (Kaufmann.timothy@mayo.edu).

See the article by Bangalore Yogananda et al. in this issue, pp. 402–411.

It is old news that the artificial intelligence (AI) revolution is well under way, not only to guide financial investments and to produce self-driving cars, but also to drive change within neuro-oncology. Perhaps in medical imaging more than any other medical field, there is great interest in what can be done with AI. We would like to know where this will all lead, and both excitement and angst hum along underneath it all. What does seem to be a foregone conclusion is that large changes will come and that they will be disruptive. We must embrace this if AI fulfills its promise, but along that path we must also carefully review how our AI tools are validated.

Among the AI techniques used with diagnostic imaging, deep learning models are now outperforming traditional machine learning ones.[1,2] While supervised machine learning utilizes features within images which are identified by humans as being relevant to the prediction of molecular markers or other important information in its classification tasks, the deep learning computer discovers all on its own those features or signals that are most relevant. Yogananda and colleagues[3] achieve 97% accuracy in predicting isocitrate dehydrogenase (IDH) mutation status in infiltrative gliomas using only T2-weighted (T2w) imaging and a fast deep learning algorithm. This compares to accuracies on the order of 80–90% when using human-selected features with traditional machine learning.[4] Previous deep learning attempts have raised this accuracy to 89–94%,[4] though these have required more preprocessing work than Yogananda's procedure. Both the robust accuracy of this current work and technical aspects of how it was achieved are important.

First, Yogananda et al achieved essentially identical accuracy with a deep learning algorithm which employed only T2w imaging as with an algorithm using a combination of T2w, T2w fluid attenuated inversion recovery, and post-gadolinium T1-weighted imaging. This is significant because T2w imaging is easily, quickly (2–3 minutes), and reliably acquired, and the use of T2w imaging from multiple different institutions and scanner platforms across The Cancer Genome Atlas (TCGA) suggests that it should be robust and generalizable. Patient motion is much less of a problem if only one quick scan is needed. The contrast enhancement of gliomas has always been considered an important imaging feature with diagnostic and prognostic implications, but here we see that it is not necessary for outstanding predictions with deep learning, given the amount of information inherent in T2w imaging alone. It is also possible that the other image series were more variable in how they were acquired and thus introduced more noise than signal.

Second, Yogananda et al's approach requires only minimal image preprocessing and does not require tedious manual or deep learning–derived tumor segmentation. Once trained, the authors' algorithm with a modern GPU (graphics processing unit) analyzes a patient's scan in only 3 minutes. Hence, their method is more practicable than some others' published methods that require a human to outline the tumor, which delays the process and may introduce greater variability. It seems that we are now close to a fast and accurate tool for determining preoperative IDH mutation status in human gliomas.

Why is preoperative IDH mutation status prediction important? Some gliomas are in locations difficult to adequately biopsy. Small sample size can also foil IDH analysis; only 35% of biopsy samples in a TCGA report contained sufficient tissue or nucleic acid quality for molecular characterization.[5] Knowing IDH status preoperatively could influence the surgeon's aggressiveness in resection because patient survival may correlate with extent of non-enhancing tumor resection in IDH-mutated tumors.[6] IDH mutation status within a tumor is also heterogeneous. Immunohistochemistry for IDH mutation shows heterogeneous cell staining[7] and fails to detect 15% of IDH-mutated gliomas.[8] The more frequently used next-generation sequencing methods have substantial error rates relative to Sanger-sequenced DNA methods.[9] It is also possible that the high failure rate in obtaining IDH mutation results relates to sample bias, if tumors that grow and necrose in a manner amenable to IDH determination are different from those that are challenging to analyze. An imaging approach to prediction of IDH mutation status would give a more global view of a tumor than would a selected biopsy, and it raises the question of the reference standard.

Is there anything to be skeptical about with regard to deep learning for brain tumor molecular or genetic characterization? First, because deep learning algorithms require relatively large training sets in order to be generalizable (ie, not overly specific to the features within their particular training sets), the authors emphasize that their algorithm needs to be tested in another patient population. This same requirement for large amounts of patient data for deep learning algorithm training is, parenthetically, a limitation to its use for rarer diseases such as pediatric primary brain tumors. Second, the regulatory and reimbursement issues surrounding the clinical use of deep learning algorithms are daunting and well beyond the scope of this editorial. Third, deep learning approaches have been criticized as being "black box," as humans are challenged in extracting and understanding the features that the computer identifies as being important in making predictions or classifications.[2,10] A lack of understanding of these features also undermines the confidence of clinicians in using deep learning and of regulators in approving its algorithms. While deep learning could be used to make great predictions as to prognosis and response to therapies, perhaps even beyond what IDH mutation status tells us, it would be very hard to learn mechanisms underlying these predictions through deep learning.

The future of neuro-oncological imaging will include deep learning, both for the prediction of tumors' molecular and genetic characteristics and for other applications such as differentiating treatment effect from progressive tumor in the posttreatment setting. Deep learning will change our practice, and hurdles to its clinical implementation are slowly but surely being overcome. The current manuscript by Yogananda et al describes important steps toward this goal.

This text is the sole product of the authors and no third party had input or gave support to its writing.

References

1. Shaver MM, Kohanteb PA, Chiou C, et al. Optimizing neuro-oncology imaging: a review of deep learning approaches for glioma imaging. *Cancers (Basel).* 2019;11(6):pii: E829.
2. Korfiatis P, Erickson B. Deep learning can see the unseeable: predicting molecular markers from MRI of brain gliomas. *Clin Radiol.* 2019;74(5):367–373.
3. Bangalore Yogananda CG, Shah BR, Vejdani-Jahromi M, et al. A novel fully automated MRI-based deep learning method for classification of IDH mutation status in brain gliomas. *Neuro Oncol.* 2020;22(3):402–411.
4. Lotan E, Jain R, Razavian N, Fatterpekar GM, Lui YW. State of the art: machine learning applications in glioma imaging. *AJR Am J Roentgenol.* 2019;212(1):26–37.
5. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* 2008;455(7216):1061–1068.
6. Beiko J, Suki D, Hess KR, et al. IDH1 mutant malignant astrocytomas are more amenable to surgical resection and have a survival benefit associated with maximal surgical resection. *Neuro Oncol.* 2014;16(1):81–91.
7. Agarwal S, Sharma MC, Jha P, et al. Comparative study of IDH1 mutations in gliomas by immunohistochemistry and DNA sequencing. *Neuro Oncol.* 2013;15(6):718–726.
8. Cryan JB, Haidar S, Ramkissoon LA, et al. Clinical multiplexed exome sequencing distinguishes adult oligodendroglial neoplasms from astrocytic and mixed lineage gliomas. *Oncotarget.* 2014;5(18):8083–8092.
9. Wall JD, Tang LF, Zerbe B, et al. Estimating genotype error rates from high-coverage next-generation sequence data. *Genome Res.* 2014;24(11):1734–1739.
10. Philbrick KA, Yoshida K, Inoue D, et al. What does deep learning see? Insights from a classifier trained to predict contrast enhancement phase from CT images. *AJR Am J Roentgenol.* 2018;211(6):1184–1193.